ELSEVIER

Discrete Mathematics 137 (1995) 357–360

DISCRETE MATHEMATICS

Note

The (Generalized) Secretary's Packet Problem and the Bell numbers

Knut Dale*, Ivar Skau

Telemark College, N-3800 Bø, Norway

Received 22 October 1990; revised 18 September 1992

Abstract

The *derangement* problem for a 'two-dimensional' version of the 'problème des rencontres' is formulated and solved by Fisk (1988). In this follow-up we find the complete probability distribution of the 'hit' variable, with which the Bell numbers appear to be connected, as the kth binomial moments multiplied by $k!$.

1. Introduction

The setting for Fisk's generalization is the following: On each of n days, p letters and envelopes are prepared without inserting the letters into the adjacent correctly addressed envelopes (all envelopes have different addresses). Then the n stacks of letters are randomly permuted, followed by a random (internal) permutation of each of the letter stacks. (The n stacks of envelopes are not moved.) Now the letters are inserted into the adjacent envelopes. Let X — the hit variable — be the number of correctly addressed letters. The derangement probability, i.e. $P_{n,p}(X=0)$, and its limit as $n, p \to \infty$, found by Fisk, are given by (3)–(4) and (7) below for $t=0$. In fact, using a general probabilistic result, it is rather easier to find the *complete* distribution of X, given by its generating function

$$F_{n,p}(t) = \sum_{x=0}^{\infty} P(X=x)t^x, \tag{1}$$

than working directly with the derangement case.

* Corresponding author.

SSDI 0012-365X(93)E0136-R

2. The finite case

Let K denote the number of letter stacks not moved, and Y be the number of letters in a p-letter stack not moved when its letters are internally permuted. Then we have

$$X = Y_1 + Y_2 + \cdots + Y_K, \tag{2}$$

i.e. X is a sum of a random number, K, of independent and identically distributed variables.

Hence it follows (see [4, pp. 12–13]) that

$$F_{n,p} = f_n \circ f_p, \tag{3}$$

where f_n and f_p are the generating functions for K and Y. Now, since both K and Y are hit variables for the classical case, we have the well-known formulas (see e.g. [5, Ch. 3])

$$f_m(t) = \sum_{k=0}^{m} \frac{D_{m-k}}{k!(m-k)!} t^k = \sum_{k=0}^{m} \frac{(t-1)^k}{k!} \quad (m = n, p), \tag{4}$$

where D_m is the number of derangements of m symbols.

Since $E(Y) = E(K) = 1$ and $\mathrm{Var}(Y) = \mathrm{Var}(K) = 1$, we note that

$$E(X) = E(K)E(Y) = 1,$$

$$\mathrm{Var}(X) = E(K)\,\mathrm{Var}(Y) + E(Y)^2\,\mathrm{Var}(K) = 2.$$

Now, since

$$F'_{n,p}(t) = F_{n-1,p}(t) f_{p-1}(t) = \sum_{k=0}^{\infty} P_{n-1,p}(k) t^k \sum_{r=0}^{p-1} \frac{D_{p-1-r}}{r!(p-1-r)!} t^r, \tag{5}$$

we end up with the recurrence formula

$$P_{n,p}(k+1) := P_{n,p}(X = k+1) = \frac{1}{k+1} \sum_{r=0}^{m} \frac{D_{p-1-r}}{r!(p-1-r)!} P_{n-1,p}(k-r), \tag{6}$$

where $m = \min(p-1, k)$. Recall that for $n = 1$ or $k = 0$ the probability $P_{n,p}(k)$ is known.

If we apply the formula for derivatives of composite functions to (3), we may obtain an explicit formula for $P_{n,k}(k)$ as a finite expression in $\{D_n\}$ and the (exponential) partial Bell polynomials (see [1, p. 139]). However, (6) should be used for calculations.

3. The limit distribution

From (3) and the 'continuity theorem' (see [2, p. 280]) we get the generating function for the limit distribution when $n, p \to \infty$:

$$F(t) = \sum_{k=0}^{\infty} P_k t^k = f(f(t)) = \exp(e^{t-1} - 1), \tag{7}$$

where $f=\lim f_n$. Recalling that the exponential generating function for the Bell numbers $\{B_k\}$ is

$$\exp(e^t-1)=\sum_{k=0}^{\infty} B_k \frac{t^k}{k!}, \tag{8}$$

we see that the numbers $B_n/n!$ are the binomial moments of this limit distribution (see [5, pp. 30–32]):

$$\frac{B_n}{n!}=\sum_{k=0}^{\infty}\binom{n+k}{k}P_{n+k}, \quad P_n=\frac{1}{n!}\sum_{k=0}^{\infty}\frac{(-1)^k}{k!}B_{n+k}. \tag{9}$$

To get a recurrence relation for $\{P_n\}$, let $n,p\to\infty$ in (6), or apply the same standard procedure as in (5)–(6) to $F(t)$:

$$P_{n+1}=\frac{e^{-1}}{n+1}\sum_{k=0}^{n}\frac{P_k}{(n-k)!}, \tag{10}$$

from which it may be deduced that $\{P_n\}$ is strictly monotone.

From (7) we get the following two explicit formulas:

$$P_n=\frac{e^{-1}}{n!}\sum_{k=0}^{\infty}\frac{e^{-k}}{k!}k^n=\frac{P_0}{n!}\sum_{k=0}^{n}S(n,k)e^{-k}, \tag{11}$$

where $S(n,k)$ are the Stirling numbers of the second kind. The first follows by a (doubly) power series expansion. The second follows from the formula (easily verified by induction):

$$\frac{d^n}{dt^n}[H(e^t)]=\sum_{k=0}^{n} e^{kt}H^{(k)}(e^t)S(n,k) \tag{12}$$

applied to $H(u)=\exp(ue^{-1}-1)$.

For an efficient procedure to compute P_n, let $R_n=n!P_n$. From (10) and the inverse binomial formula we get

$$\frac{R_n}{e}=\sum_{k=0}^{n}(-1)^{(n-k)}\binom{n}{k}R_{k+1}=\Delta^n R_k|_{k=1}, \tag{13}$$

i.e. we may calculate $\{R_n\}$ by constructing a table of differences following the pattern in [1, p. 212].

Calculations have shown that P_k is a good approximation to $P_{n,p}(k)$ even for moderate values of n and p ($n,p\geqslant 5$) when $k\leqslant\min\{n,p\}$.

The asymptotic behaviour of the distribution follows directly by applying Hayman's theorem as described in [6, pp. 153–155, condition A]. The result is

$$P_n\sim\frac{F(y)y^{-n}}{\sqrt{2\pi n(y+1)}} \quad (n\to\infty), \tag{14}$$

where $y=y(n)$ is the solution of the equation $ye^y=en$.

From (14) it may be deduced that

$$\frac{P_{n+1}}{P_n} \sim \frac{1}{y(n)} \quad (n \to \infty). \tag{15}$$

References

[1] L. Comtet, Advanced Combinatorics (Reidel, Dordrecht, 1974).
[2] Feller, An Introduction to Probability Theory and its Applications, Vol. I (Wiley, New York, 1957).
[3] S. Fisk, The Secretary's Packet Problem, Math. Mag. 61 (2) (1988) 103–105.
[4] Karlin and Taylor, A First Course in Stochastic Processes (Academic Press, New York, 1975).
[5] J. Riordan, An Introduction to Combinatorial Analysis (Wiley, New York, 1958).
[6] H.S. Wilf, Generating functionology (Academic Press, New York, 1990).